As computers have swept the modern world, expert systems have been applied in a wide variety of domains from finance (Cohen & Lieberman, 1983) to geoscience (Gaschnig, 1982) to medicine (Shortliffe, 1976). Reading educators, too, have been interested in their applications.

An expert system is a computer program that attempts to embody the knowledge of an expert in some specific domain. A consultation with an expert system resembles a teletype conversation. The expert system begins by asking questions about the problem to be solved. When the needed information has been gathered (input by the user), the system offers suggestions about how the problem can be solved.

Expert systems have been proposed as helpful in a variety of educational areas. Balajthy (1986) suggested that diagnostic expert systems could contribute to both reading research and practice and has proposed that educators become more familiar with their use and development. Vinsenhaler, Weinshank, Wagner, & Polin (1987) suggested development of diagnostic computer simulations for training diagnosticians. McEneaney (1987a) presented a prototype expert system to assist teachers in selecting reading materials for specific instructional objectives and suggested (1987b) other applications of expert systems in education.

The purpose of this article is to introduce educators to some basic concepts on which expert systems are based, to consider how expert systems might be productively applied in education, and to describe an experimental expert system with applications in reading diagnosis and teacher training.

**Introduction to expert systems**

Expert systems are usually intended for use by non-experts. The goal is to make expert knowledge available to nonexpert users in a manner that is convenient, easy to use, and cost effective.

One difference between expert systems and other
computer programs is the way they go about solving problems. Traditionally, programming a computer consisted of providing it with a set of steps (an algorithm) that is guaranteed to succeed if it is followed. Computers are good at solving problems with algorithms because they can follow those instructions without error and much faster than people. But not all problems can be solved in this way.

Consider, for example, a person who has just lost a contact lens. (This is an elaboration of an example described by Harel, 1987, pp. 342-343.) Solving the problem means finding the lens, but there are a variety of ways a person can try to do this. One way is a blind search—just look wherever your eyes happen to fall and keep looking until you find the lens. Another approach is a systematic search. You might begin by searching the ground immediately around your feet and gradually extend your search in concentric rings until you find the lens. A third possibility is an analytic search, applying what you know about wind, gravity, and the aerodynamics of contact lenses to predict where the errant lens has fallen.

Probably, however, if you lost a contact lens you wouldn’t use any one of these three search methods. You’d probably carry out a heuristic search, one that includes elements from several search approaches and your best guess about where the lens fell. You might, for instance, limit your search to a relatively small area since the lens cannot have gone very far. You might also limit your search to your left side if you felt something that might have been the lens brush by your left hand. Once you narrowed your search in this way, you’d carry out a systematic search within that small area.

The main problem with a heuristic search is that it can never guarantee that you’ll find the lens. Perhaps a freak gust of wind you didn’t even feel carried it outside your search area. Perhaps it was just an airborne seed that brushed by your hand and your lens is really in your pants cuff. More often than not, however, the assumptions that allow you to narrow your search are both plausible and productive since the other approaches to searching require more time or information than you have at your disposal.

Most of the problems expert systems have been developed to solve are like finding a lost contact lens—they are complex real-world problems that are not easily reduced to formulas or they require solutions under conditions of incomplete information. Under conditions like these, a heuristic approach provides a best-bet route to a solution even though a solution is not guaranteed.

Components of expert systems. Expert systems typically have three basic components: a knowledge base, a user interface, and an inference engine.

(a) The knowledge base is where the information used to solve problems is stored. At the beginning of a consultation, the knowledge base will be limited to the general information that an expert in that domain would be expected to know.

General information often takes the form of rules: “If such and such is the case, then so and so follows.” But solving specific problems requires more than general rules. It also requires problem-specific information and that is the purpose of the initial questioning of the user by the system—to add problem-specific information to the knowledge base.

The goal of the expert system is to solve the problem by integrating the general domain-specific knowledge that is built into the system with the problem-specific knowledge provided by the user.

(b) The user interface is the part of the program that controls the conversation between user and computer. The user interface determines whether your conversation consists of selecting items from menus, responding yes or no to questions, or filling in forms. The user interface is also responsible for the degree to which the system can explain its solutions or otherwise assist users.

(c) The inference engine is the heart of the expert system since this is the part of the program that builds the bridge between information and solutions. Two different approaches to problem solving are usually distinguished, and inference engines are accordingly characterized in two different ways, as either backward chaining or forward chaining.

A backward-chaining inference engine is employed when a solution to a problem is selected from a set of possible solutions that are known beforehand. An example of a problem that can be solved with a backward-chained approach is diagnostic classification where the number and kind of diagnostic categories are known beforehand. The problem is to
select the diagnostic categories that apply. A backward-chaining approach simply tests each of the possible solutions to see if, in fact, they are consistent with the available information.

A forward-chaining inference engine is most appropriate when the set of possible solutions cannot be specified beforehand, usually because solutions must be individually tailored for specific problems. Forward-chained problem solving is more open ended and requires that a solution be constructed rather than selected. Construction of the solution consists of identifying those aspects of the problem that are important and then combining appropriate solution components in a manner that satisfies the problem constraints.

An example of a problem typically solved with a forward-chained approach is a configuration problem, such as designing a schedule or arranging the components of a machine or electronic device. Configuration problems are not appropriately solved with a backward-chained approach because there are too many possible solutions to realistically consider each one. Instead, the forward-chained approach solves the problem a step at a time, ultimately constructing a solution from a collection of subsolutions.

**How can expert systems be applied in education?**

One expert system application that has already been developed and put to use in education focuses on the identification of children with special learning problems. Ferrara and his colleagues at Utah State University have developed a variety of expert systems to aid educators in the diagnosis of behavior disorders (Ferrara, Baer, Althouse, & Reavis, 1988), learning disabilities (Ferrara, Hofmeister, Althouse, & Likins, 1988), and developmental handicaps (Giere, Williams, & Ferrara, 1988).

Another expert system application that has already been developed as a prototype is RIMES (Reading Instructional Materials Expert System, McEnaney, 1987a), a program designed to assist in the selection of materials for reading instruction. Such a system might be used either by individual teachers or by school systems as a preliminary screening device in the often complicated and time consuming process of selecting basal reader systems.

A third application focuses on teacher education issues and applies expert systems as teaching tools. Thornburg, Baer, Ferrara, & Althouse (1990) have described how the expert systems developed at Utah State have been used to teach the skills and concepts needed to recognize specific handicapping conditions. A related application (McEnaney, 1991), originally designed to diagnose specific reading problems (a predecessor of the diagnostic system described below), has also been used to model diagnostic reasoning for students and to provide a simulation environment for students to practice diagnostic decision making.

**Expert systems in reading diagnosis and teacher training.** Expert systems have the potential to make a number of important contributions to diagnostic theory and practice in reading. Diagnostic theory stands to benefit from the systematic examination and explicit definition of rules and diagnostic theory required in developing a computer-based diagnostic aid. Diagnostic practice stands to benefit from the systematization of practice, the support of an explicit framework for diagnosis, and from the power and convenience of computers in information management and storage.

Although reading diagnosticians as a group appear to agree on what should be included in a diagnosis (assessments of reading potential and skills, and an exploration of causal factors), individual diagnoses tend to “include a large number of one-time-only statements” that “are not reliably linked with treatment” (Vinsonhaler, Weinshank, Wagner, & Polin, 1983, p. 160). In other words, even though reading diagnosticians seem to be in agreement concerning what kind of information is called for in making a diagnosis, that information does not appear to be applied systematically, either across diagnosticians or even across nearly identical cases by a single diagnostican.

The development of a computerized diagnostic expert system requires the developers to systematize their thinking about how information is related to diagnoses and remedial practice. This systematization represents the first step. It requires a more uniform conceptual language than is currently available so
that there will be less likelihood of ambiguity and greater potential for a constructive approach to diagnostic theory and practice that will allow researchers to build on or modify the work of predecessors. Building expert systems to assist in reading diagnosis may, therefore, contribute to our conceptualization of the diagnostic process itself and may even lead to new theoretical developments arising out of the expert system tools and concepts.

A second more obvious potential benefit of the systematization required to develop expert systems is that, once such a system is developed, its use ensures consistency of diagnostic practice. A diagnostician using an expert system will not overlook or forget any important information and the manner in which information is applied in determining diagnoses, and remedial strategies will be at least consistent across cases.

Very few schools have a resident diagnostic specialist. Some larger school systems have a reading center or clinic but even in this favorable circumstance the limited number of specialists often is a bottleneck in the process of diagnosing and addressing reading difficulties. An important practical benefit of computerized expert systems is that they could fill the diagnostic gap.

Imagine a computer in the teacher workroom with a library of expert system software. A teacher could come in, log on, and discuss his or her “problem” student with the system. The system would begin by asking some questions about the student, and then, following the needs and interests of the teacher, could provide diagnostic or instructional advice.

One system might offer advice about whether a student should be referred for evaluation for special services. Another might offer advice about sources of college scholarship money. A third might provide advice about grouping students for instruction or how to improve areas of weakness for individuals or the class as a whole from last year’s standardized test data.

Unlike a human expert, such a system would provide advice at a moment’s notice and virtually around the clock. It might advise, of course, that a reading clinician be consulted, but there is good reason to believe that many problems could be solved without the expense or time associated with clinical evaluation.

Computer-based diagnostic aids’ information storage and retrieval capabilities will make important contributions to reading diagnosis and remediation. Diagnostic records for individual students can be maintained and searched with little effort. Moreover, with standardized computer-based record keeping, school systems and even states could begin to develop large databases that could be used both in the evaluation and improvement of educational programs.

Enhanced access to information also means that an expert system can make prodigious amounts of information available to users. Consider, for example, a system designed to assist teachers in the remediation of reading problems. It could be integrated with a CD-ROM video system showing users videos of example lessons illustrating remedial approaches to instruction. With the data storage capability of CD-ROM, the information can be encyclopedic both in scope and depth, so that an expert system can function as a teaching/learning tool as well as a decision-making aid.

The computer must still be programmed to deliver diagnostic instruction, but the knowledge base of an expert system can, in principle, serve as the foundation for a teacher-training system. One well-known system developed in this manner was GUIDON (Clancey, 1983), an intelligent tutoring system for medical education that was based on the MYCIN (Shortliffe, 1976) expert system.

The logic of educational diagnosis and remediation. Problem solving in education consists generally of characterizing the learner and delivering appropriate instruction. In the context of reading, this means diagnosing strengths, weaknesses, interests, etc., and then planning and executing a program of instruction.

Traditionally, clinical reading diagnosis attempts to identify general global measures that indicate student performance compared to other children (i.e., norm-referenced measures) as well as specific skill assessments (criterion-referenced measures) that indicate student competence in some set of reading skills (decoding, sight recognition, etc.).

Although almost any combination of global and skill measures is possible, diagnostic assessment in
reading is not nearly as open ended as it might appear. It seems that, like in medicine, reading diagnosis involves the successive narrowing of possibilities characteristic of backward-chained reasoning. For example, in any given skill area, a student’s score is often characterized as superior, adequate, or inadequate. Diagnosis in that skill area means identifying which of these possibilities holds true.

The selection of individualized teaching methods and materials, however, is a genuinely open-ended task and requires the constructive approach of forward chaining. Individualized instruction means selecting from a variety of methods and materials, and the decisions a teacher makes about combining those materials in specific ways multiplies the number of possible lessons considerably. Planning instruction, therefore, appears to require a very different approach to problem solving than that required in diagnosis.

This suggests an important characteristic of the logic of problem solving in education: the two components of clinical problem solving that have been identified (i.e., diagnosis and prescription) appear to require two different problem solving approaches— a backward-chained approach in diagnosis and a forward-chained approach in lesson planning. Not surprisingly, this problem-solving logic has important implications for developing expert system applications.

It means that expert systems developed for classroom use should integrate a backward-chaining diagnostic system with a forward-chaining prescriptive system. Any systems that employ exclusively backward- or forward-chaining engines will distort educational decision making in undesirable ways.

Significantly, most inexpensive expert system shells provide for only one kind of inference engine (usually backward-chaining). Thus educators should be cautious in promoting or developing expert systems using commercial shells—there is reason to believe that the decision making implemented in them distorts the logic of educational problem solving.

A session with Teacher’s Aide
Teacher’s Aide (TA) is an experimental expert system that has been developed especially for the reading classroom. Specifically, it assists in the diagnosis and remediation of reading difficulties.

The current version of TA integrates standardized test scores and results from an informal reading inventory to provide users with both a general diagnosis and an assessment of student strengths and weaknesses. Data entry typically requires 15 to 20 minutes for experienced TA users.

One of the advantages of TA is that it can be operated entirely from monitor-displayed menus. In addition, TA can be operated from a natural language interface (NLI) that allows users to communicate with TA using ordinary English. Although the NLI built into TA is simple, preliminary trials suggest it works effectively because the domain of discourse is narrow.

Figure 1 depicts TA’s screen with the diagnosis pop-up menu selected (a listing of diagnostic facts is behind the pop-up). The menu bar appears at the top of the screen. The NLI is immediately below the menu bar, and the output window (where the conversation between user and system is displayed) is at the bottom. By moving the cursor across the menu bar (with the arrow keys) the user selects options from the menu bar. When an option is selected and the down arrow or return key is hit, a pop-up menu appears immediately below the menu option selected.

Several kinds of diagnosis are supported by the TA system. A directed diagnosis begins with a set of questions that are used by the system to make one or more preliminary diagnostic hypotheses. Only diagnoses that are identified as diagnostic hypotheses are explored in a directed diagnosis, whose advantage is that typical cases are diagnosed more efficiently, since the preliminary questioning eliminates unlikely diagnoses and places the remainder in order of likelihood.

A nondirected diagnosis, on the other hand, considers every diagnostic category recognized by the system and is, therefore, better suited to diagnosing atypical cases that might be ruled out in a directed diagnosis.

In addition to directed and nondirected diagnoses, the TA system supports a diagnostic check to see if a student satisfies diagnostic criteria for a specific diagnosis. A diagnostic check might be preferred by a user if there is reason to believe a specific diagnosis.
Figure 1
User interface in the Teacher’s Aide (TA) experimental expert system

<table>
<thead>
<tr>
<th>File</th>
<th>Diagnose</th>
<th>Recommend</th>
<th>Show</th>
<th>Explain</th>
<th>Move</th>
<th>Help</th>
</tr>
</thead>
<tbody>
<tr>
<td>NL1</td>
<td>Directed diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nondirected diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUT</td>
<td></td>
<td>Diagnostic check for ...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intel</td>
<td></td>
<td>• Average reader</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISTEP</td>
<td></td>
<td>• Overachiever</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More</td>
<td></td>
<td>• Underachiever</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>There</td>
<td></td>
<td>• Slow learner</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>There</td>
<td></td>
<td>• Gifted reader</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indep</td>
<td></td>
<td>• Disabled reader</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic</td>
<td></td>
<td>Miscue analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indep</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Readi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reading expectancy quotient is greater than 110.
Reading achievement is above grade placement.
IRI instructional level is above grade placement.
Three IRI levels (Ind, Inst, Frus) are identified.

Teacher’s Aide (TA) copyright © 1991 John E. McNeaney

is likely or if the only purpose of diagnosis is to see if the reader meets certain diagnostic criteria (e.g., does the student qualify for special services).

Since the diagnostic check addresses specific diagnostic criteria, it can usually be carried out in less time than a complete diagnosis. In the current version of TA, diagnostic checks are limited to general diagnostic categories (normal developmental reader, disabled reader, underachieving reader, etc.) but there is no reason more specific kinds of diagnostic checks cannot be incorporated. Skill-specific diagnostic checks could be incorporated if suitable rules were added to the system. With the addition of a rule editor, a system like TA could even allow users to customize the system.

Finally, in addition to diagnostic categorization, the TA system supports diagnostic skill assessment in reading using a combination of user questioning and assessment data. Six skill areas are addressed: sight vocabulary, phonic analysis, structural analysis, contextual analysis, comprehension, and oral reading fluency and rate.

The TA knowledge base consists of rules, facts, frames, and system statistics. The knowledge base can be examined using the Show option on the main menu that is depicted in Figure 2. The Show option submenu allows users to inspect the information the system uses in making its diagnostic assessments.

Rules are the general knowledge that the system uses in diagnosis. Facts are simple declarative statements representing knowledge gathered about the reader. Frames are organized sets of data that have specific diagnostic applications. System statistics are data TA keeps about the diagnoses it makes; they are employed by TA to set up the order in which diagnostic hypotheses are examined. (More common conditions are tested first.)

When the information needed to make a diagnosis has been entered, the system offers its diagnosis. TA can also explain its diagnosis and in a newer version currently under development will make instructional recommendations (using the Recommend menu option). A complete record of the diagnostic session (user requests and system responses) is recorded.
and, in addition, users can enter informal diagnostic notes and observations with a text editor that is part of the system.

**Diagnosing a reader: Maria.** Let’s take an example of a diagnostic session. It begins with the user entering case data for “Maria”—information such as name, an identifying number, age, years in school, etc. Case data must be taken into account in many or all of the diagnostic decisions the system makes, so these data are set aside as a special class.

Case data are an example of the frame-based approach to knowledge representation described above. Case data are entered via a case data screen that simulates a form to be filled out.

After gathering case data, TA begins its inquiries about diagnostic measures. The first (and most complex) set of questions deals with student performance on an informal reading inventory (IRI). IRI questions begin with queries about isolated word recognition and then ask for detailed information about responses to comprehension questions and about numbers and kinds of oral reading miscues.

Student performance on word lists is entered on a chart. Data requested include the grade level of the list, number of words on the list, number of words recognized instantly, number of words recognized in an untimed format, and other data concerning the number and kind of specific miscues.

After word list data have been entered, the user is queried about student performance on IRI reading passages. The passage data screen is similar to the word list screen, but in addition to the chart, each row of data is followed up with a set of questions concerning passage miscues. When IRI passage data have been entered, the system computes word recognition and comprehension percentage scores for each passage, identifies the client’s reading levels, and prompts the user to press a key to continue.

Maria’s IRI passage data and system responses are depicted in Figure 3.

When IRI data have been entered, the system inquires about standardized measures of ability and
achievement. This information is stored in cognitive and achievement frames and is used to compute reading formulas (Harris & Sipay, 1985) that are the basis of under- and overachievement diagnoses. Since TA was originally developed for use with students in Indiana, the achievement scores requested are those of the Indiana Statewide Test of Educational Progress (ISTEP).

Since the diagnostic classifier is backward chaining, gathering of data and diagnostic assessment occur simultaneously. The first step is to select a tentative diagnosis. The second step is to test each of the defining conditions that support that diagnosis to see if they hold true for the client.

One by one, each of the defining conditions is tested by asking the user for the appropriate information. When a condition is found to hold true for the client, that data is added to the current client database. If a condition fails to hold for a client, that data is added to the knowledge base and the diagnosis is eliminated from further consideration.

If every condition for a diagnosis is satisfied, the system reports a successful diagnosis and offers an explanation. If the data fail to support any diagnosis, the system reports a failure to diagnose and offers an explanation for its failure.

The data provided for Maria result in a high ability—disabled diagnosis. As a reader of superior ability, Maria would be expected to perform at a higher level than her average classmates but IRI results indicate average performance, as do her standardized test scores. Since the TA system explicitly relates achievement and ability, it successfully recognizes the discrepancy between ability and achievement, identifying her as a child experiencing reading difficulties despite her grade-level performance on the ISTEP and an instructional reading level at the fourth grade (her current placement).

The explanatory framework supporting Maria’s diagnosis and the rules it is based on are depicted in the Table. The highest level rule in the diagnosis is Rule 10: “The student appears to be a high ability-disabled reader.” Rule 10 has four conditions: (1) the IRI must be a valid measure, (2) global IRI error patterns should support a possible reading problem, (3) the student’s IRI instructional level should be at or
Levels of explanation available for Maria's diagnosis and the rules responsible for each decision point in the diagnosis

Diagnosis: The student appears to be a high ability-disabled reader. (Rules applied: 10, 100, 120, 153, 210, 300, 410, 503, 602, 700, 705)

LEVEL 1:
The student appears to be a high ability-disabled reader because the IRI appears to be a valid measure of the student's ability, and global IRI error patterns suggest a possible reading problem, and the highest IRI instructional level is at or below grade placement and other measures support concurrent diagnosis of high ability and disability.

LEVEL 2:
The IRI appears to be a valid measure of the student's ability because three IRI levels (Ind. Inst. Frus) are identified, and the student was cooperative and attentive during testing.

Global IRI error patterns suggest a possible reading problem because three or more instructional reading levels have been identified.

The highest IRI instructional level is at or below grade placement because the highest instructional level is at grade placement.

Other measures support concurrent diagnoses of high ability and disability because measures of general ability and achievement are available, and intelligence is above average, and ISTEP scores suggest a learning problem despite a high IQ.

LEVEL 3:
Measures of general ability and achievement are available because an IQ score is available, and ISTEP scores are available.

ISTEP scores suggest a learning problem despite a high IQ because reading achievement is at or below grade placement, and reading achievement is less than predicted by IQ, and reading achievement is below or comparable to age peers.

LEVEL 4:
Reading achievement is at or below grade placement because reading achievement is at grade placement.

Reading achievement is less than predicted by IQ because reading expectancy quotient is less than 90.

Reading achievement is below or comparable to age peers because reading achievement is comparable to age peers.

LEVEL 5:
Reading achievement is comparable to age peers because reading quotient is between 90 and 110.
below grade placement, and (4) other measures should support concurrent diagnoses of high ability and disability.

Each condition supporting a rule must be satisfied for a diagnosis. But conditions may themselves have further conditions, so that any one condition may lead to a string of lower level conditions. The purpose of layering rules (and conditions) is both to organize those rules into a hierarchical framework and to provide users with different levels of explanation.

Organizing rules into a hierarchical framework means a rule set can be understood, added to, or modified more easily. Moreover, when a user asks for an explanation of Maria’s diagnosis, the system can begin with the kind of higher level explanation that a human expert would usually begin with (i.e., Maria is diagnosed as a high ability—disabled reader because the IRI is valid, global IRI patterns suggest a problem, her instructional level is at or below grade placement, and other measures support concurrent diagnoses of high ability and disability).

If the user asks for further explanation of Maria’s diagnosis, however, the system will continue with successively more detailed explanations based on the lower level rules and conditions in the diagnosis (e.g., other measures support concurrent diagnoses of high ability and disability because measures of general ability and achievement are available, intelligence is above average, and ISTEP scores suggest a learning problem despite a high IQ). Explanations for diagnoses, therefore, have a layered quality that allows users to satisfy themselves with as much or as little detail as desired.

Limitations of expert systems
It is important to note two limitations of expert systems in reading diagnosis and education in general. One has to do with the limits of specific systems like TA. Another more general limitation has to do with the appropriate use of expert systems in educational decision making.

System-specific limitations. Specific expert systems inevitably have limitations due to the assumptions they are based on and their capacity to represent and manipulate information. Teacher’s Aide, for example, makes certain assumptions at present about the utility of standardized measures of achievement and skill analysis for the purposes of diagnosis and remediation.

One reason TA takes an essentially skills-based approach to diagnosis is that such an approach is well established in reading education and as a result is more easily translated into a diagnostic logic and a set of rules. Other system-specific limitations of TA include the deliberate exclusion (for simplicity’s sake) of arguably relevant dimensions of reading such as reader interests, silent and oral reading differences, and flexibility depending upon task and text type. It also presupposes a human expert who can administer an IRI.

- The good news about system-specific limitations of expert systems, however, is that these limitations can often be eliminated with suitable modifications or additions to the system. For example, there is nothing about expert system technology that requires a skills-based approach to diagnosis. Expert systems can be developed that adopt a more holistic or process-oriented approach to assessment.

- A process-oriented assessment system will, however, differ in a number of ways from TA. One difference is that whereas TA is designed to be used when assessment data have already been collected, a process-oriented assessment system would involve numerous consultations across time and would probably be designed to assist teachers in the collection of data as well as in the recognition and interpretation of diagnostic patterns. A process-oriented system might include observational checklists, data from running records of oral reading across time (Clay, 1985; Johnston, 1992), and informal rating scales for writing samples.

- Current technologies would allow student writing, drawing, and other kinds of work to be optically scanned and stored in computer files for analysis and viewing at a later time. Even aspects of reading such as fluency and expression could be recorded and analyzed in computer systems, with the capacity to manage audio and video input (multimedia), a capacity well within current technologies.

- A second difference is that, although a process-oriented assessment system is still based on a set of rules, those rules will generally focus on relating data that has been entered to further ongoing assessment and instructional practices rather than diagnostic
categories. Such a process-oriented system might present a teacher with a chart depicting each student's status in the ongoing program of assessment along with other recommended assessment procedures and instructional practices. This increased emphasis on instruction also suggests that a process-oriented system could lead very naturally to a computer-based management system for whole language instruction.

In fact, any approach that can be described in sufficient detail can provide a basis for a reading assessment expert system. But the computer adage "garbage in—garbage out" applies to expert systems just as it does to more traditional programs: an expert system is only as good as the assessment framework on which it is based.

A general limitation. A second more general limitation of expert systems in education arises from their potential for misuse or misinterpretation. This kind of technology may induce some diagnosticians or teachers to abdicate responsibility for decision making.

One of the reasons for incorporating the capacity to explain diagnoses is that users can evaluate both the information and the logic that leads to them. System diagnoses in TA are reported in a manner that encourages users to view them as suggestions rather than immutable truths, but there is always the danger that users will grant this kind of technology an authority that is unwarranted.

Professional ethics demand that teachers (rather than machines) make decisions. The role of the technology should be to support professional decision making, not to supplant it. Unfortunately, the problem of appropriate use cannot be solved in the straightforward manner system-specific problems usually can. The key to the solution isn't technological, it's pedagogical. Teachers need to be trained in the use of this technology and its limits.

Their limitations notwithstanding, expert systems like the one described here appear to offer educators a powerful new approach to conceptualizing and supporting classroom decision making. Expert systems have been successfully developed and applied in a variety of settings and there is good reason to believe educators can apply this technology with good results in reading education.

References
Beyond JR: Research from elsewhere
Jeanne Shay Schumm

Content area textbooks: How tough are they?

Content area textbooks are tough. This is even more true for students with special needs such as mainstreamed special education students and those with limited proficiency in English. But just how difficult are textbooks?

Diane Kinder, Bill Bursuck, and Michael Epstein of the Educational Research and Services Center and Northern Illinois University have examined widely used social studies textbooks to see just how tough they really are. The researchers selected 10 American history textbooks published since 1985. To ensure uniformity of content, only the chapter on the post-World-War-II era from each book served as the unit of analysis. Analyses included examination of readability level, global cohesion (text organization), local cohesion (pronoun references), location and type of questions included, and vocabulary density.

The readability of the 10 textbooks varied considerably, from grade levels 9 to 15 (the mean was 10.9). The books also varied in respect to clarity of pronoun references and number and type of questions included. Only one textbook included questions at the beginning of the chapter to help activate prior knowledge. All of the chapters analyzed included subheadings to promote global cohesion and most offered introductions and summaries, although the quality of the introductions and summaries was questionable in some texts.

One area of consistency among the 10 textbooks was introduction of technical vocabulary. All introduced approximately one new word per page and all highlighted key vocabulary.

The variability among the textbooks prompted the authors to recommend that teachers closely scrutinize textbooks before adoption. While this is ideal, teachers do not always have the luxury to choose their own text. To the analysts' suggestion I like to add that teachers might consider an analysis of their assigned textbook—particularly an analysis of text features that might enhance comprehension (Armbruster & Anderson, 1981; Singer, 1986). Such an analysis can help sensit化 teachers to weaknesses in the textbook that might pose problems for their students.